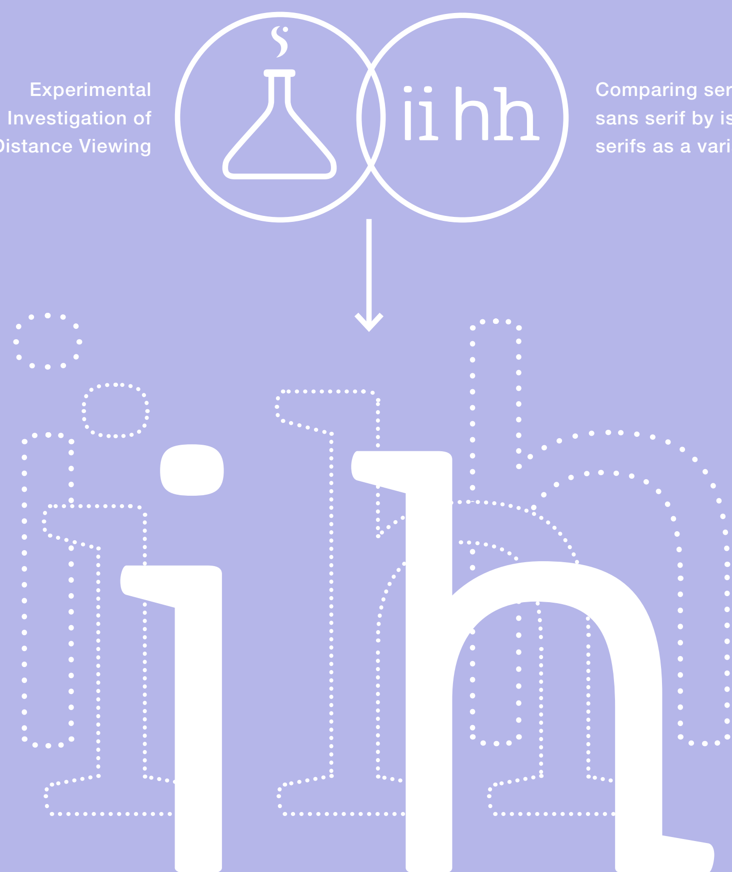
The influence of serifs on 'h' and 'i': useful knowledge from design-led scientific research

47.3



Experimental Investigation of Distance Viewing

Comparing serif and sans serif by isolating serifs as a variable

RESULT

Letters with serifs on the vertical extremes are more legible than the same sans serif letters, while lower case serif letters "i" and "h"are easily confused with serif letters "l" and "b".

BEIER, DYSON
The Influence of Serifs on "h" and "i": Useful Knowledge from Design-led Scientific Research p. 69

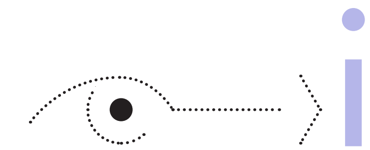
ISSN 0022-2224

47.3 VISIBLE LANGUAGE

47.3

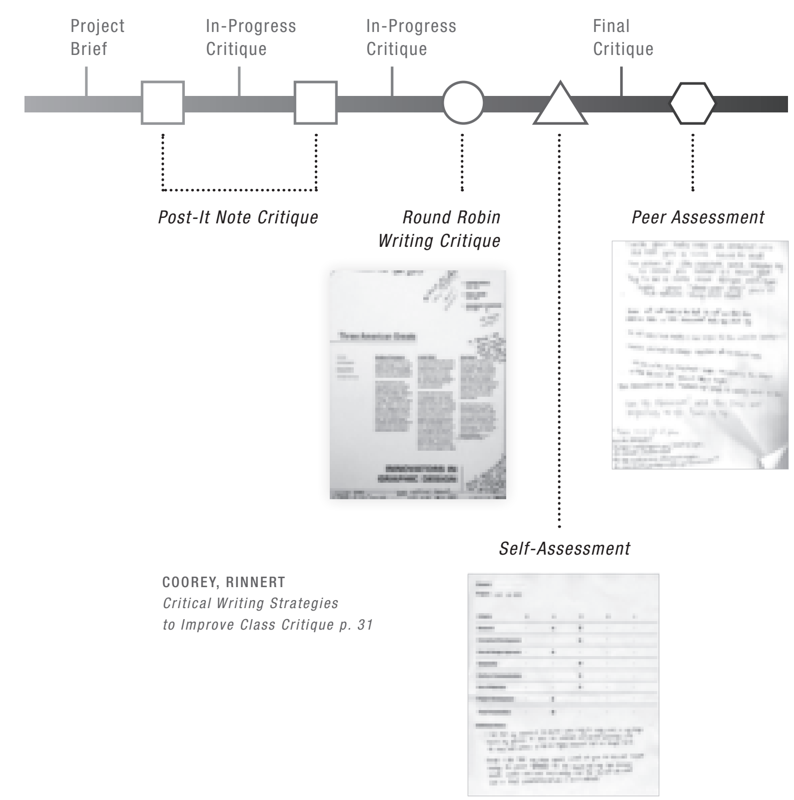
VISIBLE LANGUAGE

PMS 2705 PURPLE BLACK



THE JOURNAL OF VISUAL COMMUNICATION RESEARCH

PROJECT TIMELINE



COOREY, RINNERT
Critical Writing Strategies to Improve Class Critique p. 31

Before there was reading there was seeing. Visible Language has been concerned with ideas that help define the unique role and properties of visual communication. A basic premise of the journal has been that created visual form is an autonomous system of expression which must be defined and explored on its own terms. Today more than ever people navigate the world and probe life's meaning through visual language. This journal is devoted to enhancing people's experience through the advancement of research and practice of visual communication.

If you are involved in creating or understanding visual communication in any field, we invite your participation in Visible Language. While our scope is broad, our disciplinary application is primarily design. Because sensory experience is foundational in design, research in design is often research in the experience of visual form: how it is made, why it is beautiful, how it functions to help people form meaning. Research from many disciplines sheds light on this experience: neuroscience, cognition, perception, psychology, education, communication, informatics, computer science, library science, linguistics. We welcome articles from these disciplines and more.

Published continuously since 1967, Visible Language maintains its policy of having no formal editorial affiliation with any professional organization — this requires the continuing, active cooperation of key investigators and practitioners in all of the disciplines that impinge on the journal's mission as stated above.

WEBSITE

visiblelanguagejournal.com

POSTMASTER

Send address changes to:

Sheri Cottingim
Office of Business Affairs
College of Design, Architecture, Art, and Planning
University of Cincinnati
PO Box 210016
Cincinnati, OH 45221-0016
sheri.cottingim@uc.edu

Published tri-annually in January, May and October

BLACK

SUBSCRIPTION RATES

United States	Individual	Institutional
1 year	$35.00	$65.00
2 year	$65.00	$124.00
3 year	$90.00	$183.00
Canadian*	Individual	Institutional
1 year	$44.00	$74.00
2 year	$83.00	$142.00
3 year	$117.00	$210.00
Foreign**	Individual	Institutional
1 year	$56.00	$86.00
2 year	$107.00	$166.00
3 year	$153.00	$246.00

Prepayment is required. Make checks payable to University of Cincinnati Visible Language in U.S. currency only. Foreign banks need a U.S. correspondent bank.

Pay for your subscription onine at daap.uc.edu/CCtest.html

* Canadian subscriptions include additional postage ($9.00 per year).
**Foreign subscriptions include additional postage ($21.00 per year).

ISSN 0022-2224
Published continuously since 1967.
Index included in last issue of volume year.

BACK COPIES

A limited number of nearly all back numbers is available. The journal website at http://visiblelanguagejournal.com is searchable and lists all issues, contents and abstracts.

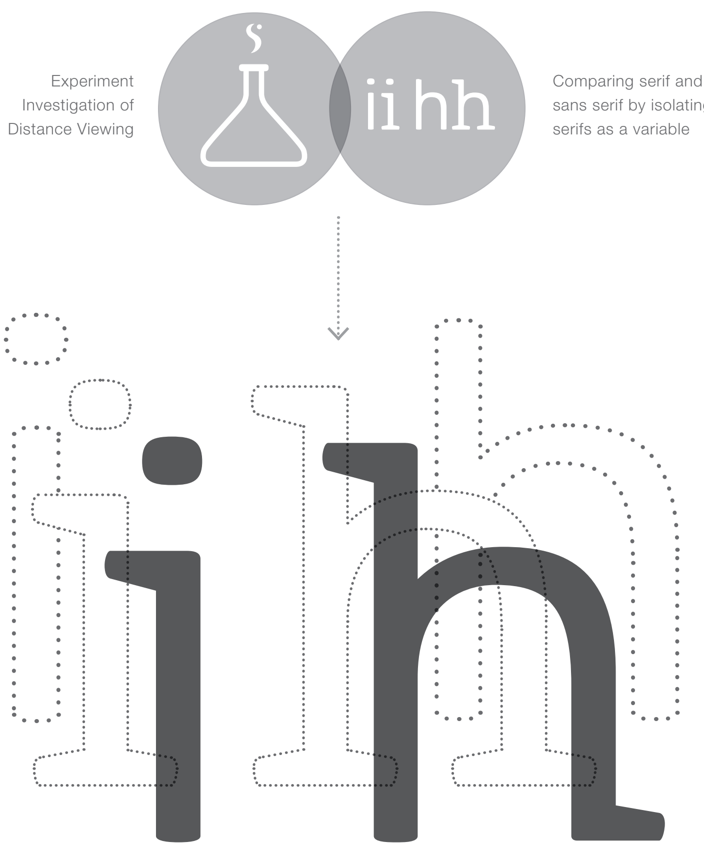
COPYRIGHT INFORMATION

VISIBLE LANGUAGE

THE JOURNAL OF VISUAL COMMUNICATION RESEARCH

47.3

47.3



RESULT

Letters with serifs on the vertical extremes are more legible than the same sans serif letters, while lower case serif letters "i" and "h"are easily confused with serif letters "l" and "b".

04 The Influence of Serifs on 'h' and 'i': Useful Knowledge from Design-led Scientific Research

DR. SOFIE BEIER, DR. MARY C DYSON

ABSTRACT

The typographical naivety of much scientific legibility research has caused designers to question the value of the research and the results. Examining the reasons underlying this questioning, the paper discusses the importance of designers being more accepting of scientific findings, and why legibility investigations have value. To demonstrate how typographic knowledge can be incorporated into the design of studies to increase their validity, the paper reports on a new investigation into the role of serifs when viewed at a distance. The experiment looks into the identification of the lowercase letters ‘j’, ‘i’, ‘l’, ‘b’, ‘h’, ‘n’, ‘u’, and ‘a’ in isolation. All of the letters originate in the same typeface and are presented in one version with serifs and one version without serifs. Although the experiment found no overall legibility difference between the sans serif and the serif versions, the study showed that letters with serifs placed on the vertical extremes were more legible at a distance than the same letters in a sans serif. These findings can therefore provide specific guidance on the design of individual letters and demonstrate the product of collaboration between designer and scientist on the planning, implementation, and analysis of the study.

INTRODUCTION

› A review of the published material on experimental legibility research demonstrates that a lack of typographical understanding among many researchers with a background in science can make these findings irrelevant to designers. This paper will offer an account of how this lack of understanding is manifested in the methodological detail of the studies, and suggest a direction for the field of legibility research. Through an example of new empirical research following the recommended approach, the paper shows how experimental investigations can inform the work of practicing designers.

According to Nigel Cross (2001) the discussion of science and its relation to design has a 40-year circle, first appearing in the 1920s, and then again in the 1960s. The impact of this relationship has been thoroughly covered by a number of authors (e.g. Cross, 2001; Heylighen et al. 2009).

A scientific work process is often described as being driven by a search for evidence, where assumptions are tested aiming at some form of answer. The skill of designing, on the other hand, is often described as a creative operation that is strengthened by a design process acquired by craftsmen through practice. Harold Nelson and Erik Stoltenman (2003) see the design approach as having two major cognitive components: one is that of imagination, intuition, emotion and instinctive thinking; the other is that of reasoning containing both synthetic and analytic thinking. According to the authors, analytic thinking — here similar to science — is a reductive way of thinking where you separate parts for investigation. Yet, only when the elements are brought back into a context that recognizes the complexity of the matter, will the analytic thinking create productive knowledge.

Although the underlying bases of the work process of scientists and designers have their differences, the two approaches can benefit significantly from the contributions of the other. Receptiveness from academic researchers towards incorporating design expertise in the development of test material in legibility investigations can help validate the research findings; the investigation will not only inform the discipline of the researcher, but also create findings that can be usable for the practicing designer. If designers on their part are more open to scientific findings, such research can provide an informed influence on their work. With greater awareness of the empirical work process, designers can further benefit from implementing this process in their designing.

In recent years, a growing interest in strengthening the connection between the two different ways of working is emerging, with various research institutions encouraging interdisciplinary work.

When the field of typeface legibility research functions at its best, it contains elements of both science and design.

The scientific approach contributes with controlled test methods and with the analysis of the data, while the design approach contributes with the creation of relevant material for testing.

THE PROBLEMS OF THE TYPEFACE LEGIBILITY FIELD

A number of experimental legibility studies comparing different typefaces were carried out in the first part of the 20th century (for English-language ones see: Tinker 1964; Luckiesh & Moss 1942; Pyke 1926; Roethlein 1912). In recent times, a large number of researchers have also published papers related to legibility (for examples see: Fiset et al. 2008; Sheedy et al. 2008; Fox et al. 2007). These studies are published in academic journals, which are often not read by the practising designer. Even if they were read, the way the material is presented often makes it difficult for practitioners to understand the findings and translate them into usable knowledge. In 1969, Merald Wrolstad, the editor of the *Journal of Typographic Research* (now *Visible Language*), explained the lack of interest in research into legibility among designers. She pointed out that since research projects tend to be motivated by the researcher's area of interest, and because many legibility researchers are engineers, psychologists, or reading specialists, the research will focus on variables of interest to these disciplines and not on the various aspects of letter-forms.

An argument voiced by a number of designers through the online typography community Typophile (2008a; 2008b) is that we must fully understand the underlying principle of legibility before studying it. This argument fails to recognize that understanding is acquired through research, and building on existing knowledge. The claim made in these threads is that no test method that includes typefaces as test material will be able to provide any useful data, due to the difficulties in isolating a single typographic stylistic feature for investigation.

Hence they argue that researchers do not understand the principle of legibility because they use typefaces in their studies that vary in many ways.

When talking about layout, readability-related variables such as line length, word spacing, and leading tend to interact, and therefore, designers may argue that it is inappropriate to isolate only one variable for experimental investigation. The argument being that a designer would always make adjustments, for example, to leading if changing line length. However, the psychologist's approach to testing line length (e.g. Dyson & Kipping 1998) isolates this one variable to ensure that the results can be attributed to this variable.

On the following pages, we argue that if the focus of the investigation is the features of a typeface, it is possible to isolate one stylistic feature for investigation.

But the material (i.e. different styles of typeface) needs to be designed with knowledge of how the different features of typefaces interact.

A lack of such insight can be observed in a number of the empirical studies into typeface legibility that make comparisons of typefaces. These typefaces tend to vary in many ways, which makes it impossible to identify which of the different features affect performance. The investigations lack sufficient understanding of what is in fact under study (for examples see: Soleimani & Mohammadi 2012; Mansfield et al. 1996; Lange et al. 1993). In the following, we will discuss three of the potential problems often seen in studies that compare typefaces of different style.

The first problem concerns the appearing size of a typeface, which varies dramatically according to the size of the x-height of the character (*figure 1*). The outcome is two typefaces of the same point size that appear to have very different actual sizes. The matter is generally understood among most graphic designers, and is a common topic in the main educational literature on typographic layouts. However, some academic researchers are either unaware of the issue, or they have chosen to ignore it due to technical limitations in scaling the fonts for in-between sizes.

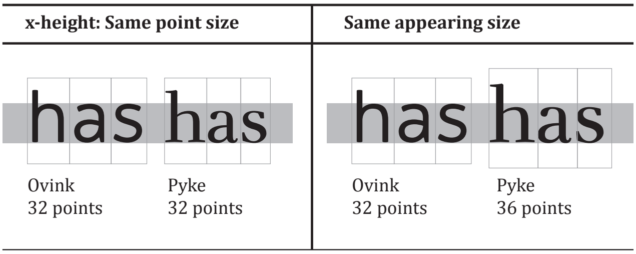


FIGURE 1 *The x-height of the typeface Pyke is considerably smaller than the x-height of the typeface Ovink. The typefaces to the left are both set in 32 point; to the right the same typefaces are adjusted to have a similar appearing size. This results in different point sizes, and ascenders and descenders having different lengths.*

To make matters more complicated, it is fairly common that the type size on the body varies considerably between typefaces — a variation that has existed since the early printers. In such cases, one typeface can have a letter height of 8 millimetres when set in 32 points, while another typeface in the same point size will have a letter height of 6 millimetres (*figure 2*). Equating point size is consequently not an appropriate means of comparing different typefaces.

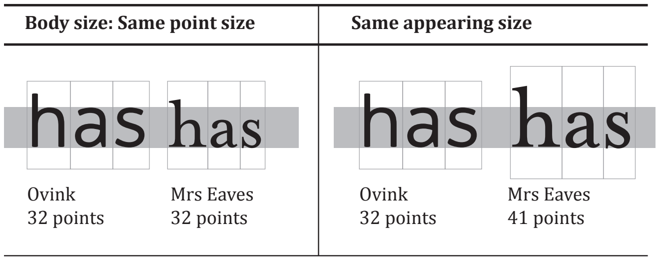


FIGURE 2 *The typeface Mrs Eaves is set at an unusually small size on the body. In this case the problem of equivalence is not only related to the x-height but to the extending parts of the letter as well. To the left are the typefaces Ovink and Mrs Eaves both set in 32 points; to the right the same typefaces are adjusted to have a similar appearing size. This results in different point sizes, and ascenders and descenders having different lengths.*

A second potential problem with test material that compares different typefaces is that typefaces tend to vary in both weight and width. There is a risk of drawing incorrect conclusions if the researcher is not aware of the possible implications of these variables. The fact that they do influence legibility is demonstrated in an experimental investigation by Barbara Elizabeth Roethlein (1912). Looking at the relative distance legibility of a number of different styles of the typeface Cheltenham, the study found that both typeface weight and width influence performance when text is read at a distance. This is an early example of a controlled investigation where the differences in stylistic features are kept to a minimum.

The raw data presented in the paper shows that the Bold style could be read at a greater distance than both the Regular (Ordinary) and the Bold Condensed style. When comparing the data for the Regular and Bold styles, we can be fairly sure that the difference in performance originates in the difference in weight. Similarly, we can be fairly sure that the difference in performance between the Bold and the Bold Condensed styles originates in the difference in width.

The results of Roethlein's study mean that if two different typefaces were compared at a distance and these happened to vary in weight and width, as well as other stylistic features, a style with heavier weight and/or greater width is likely to be read more easily than a style of lighter weight and/or narrower width. If the effect of typeface weight or width is not taken into account, the results may be attributed to other characteristics of the typeface.

A third possible problem in test materials is related to stroke contrast. Many designers know that under difficult reading conditions, the hairline strokes in high contrast typefaces tend to bleed out, and so will result in letter shapes of poor definition. If stroke contrast is not the variable of interest in a study, this too may be a confounding variable and lead to invalid conclusions.

Ironically, in the context of typefaces, it is the scientists and not the designers who fail to isolate variables for testing.

Due to the (understandable) lack of typographical knowledge among scientists, these researchers tend to view x-height, weight, width, and stroke contrast in a more holistic fashion, failing to identify these features as variables.

Consequently, a substantial number of earlier legibility investigations fail to provide designers with findings that are relevant to their work. But although the field of legibility research has its problems, typographical knowledge has informed the research hypotheses of a number of recent studies (Dyson 2011; Bias et al. 2010; Beier & Larson 2010; Larson & Sheedy 2008; Waller 2007; Morris et al. 2002). As a further example of this type of study, the present investigation sets out to explore the dogma often found among graphic designers that serif typefaces are more legible than sans serif typefaces. The aim is to identify the influence serifs have on the individual letter. The two authors have different educational backgrounds, one in graphic design and one in experimental psychology. The complementary disciplines enable collaboration on the design of the study to produce a scientifically valid methodology that addresses a question that has validity among designers.

THE ROLE OF THE SERIF

In the design community, it is often stated that when setting continuous text, a serif typeface should be used, as it is believed that sans serif typefaces are less legible. According to John D. Berry (2006), this rule of thumb is motivated by the idea that sans serif typefaces — in contrast to Old Style serif typefaces — are more mechanical and lifeless in their appearance. Type designer Gerard Unger (2007) has spoken in favour of serifs, as he argues that serifs on ascenders and descenders may help ease reading in the parafoveal view by emphasising these specific parts of the letters. Type designer Adrian Frutiger (n.d.) has a similar argument, suggesting that a stroke with no fortified endings will leave the observer with the feeling that it could flow on forever. He finds that serifs, like the base and the top of architectural columns, give emphasis to the terminations of the column. Another central matter often referred to when speaking in favour of the serifs, is the role they play in facilitating a higher differentiation between letters. According to designer Erik Spiekermann (2007), serifs helps to avoid confusions between the lowercase letters ‘l’, ‘i’, the uppercase ‘I’, and the digit ‘1’ when these letters are in isolation.

However, the majority of reading rate experiments that have compared the legibility of sans serif and serif typefaces have found the performance of the two to be similar. In a critical review of 72 studies that compare different typefaces, Ole Lund (1999) found no valid conclusion in favour of either serifs or sans serif typefaces. However, an experiment carried out by mathematician Robert A. Morris with vision scientist colleagues (2002) has shown that a serif version of the typeface Lucida, of 40 pixels x-height presented at a 4 meter distance, slows down reading compared to a version of Lucida without serifs and with slightly less stroke contrast. The difference between this study and the many it succeeded is that one of Morris’s co-authors is a type designer (Charles Bigelow).

By applying an interdisciplinary approach between fields of science and design, the research produces more rigorous and relevant findings than when the work is informed by one discipline alone.

The findings by Morris and colleagues indicate that serifs are not always either good or bad. An investigation by John Harris (1973) may explain this ambiguity, as his results suggest that serifs seem to reduce the legibility of certain letters, and improve others. In a study of the sans serif typefaces Univers 689 and Gill Sans Medium, and of the serif typeface Baskerville 169, Harris applied the methodology of briefly presenting a single letter four degrees off the centre of vision, and found that the confusions of ‘b’ > ‘h’, ‘a’ > ‘n’, ‘n’ > ‘u’, and 'u' > 'n' were significantly higher in the serif typeface compared to the two sans serif typefaces (*figure 3*).

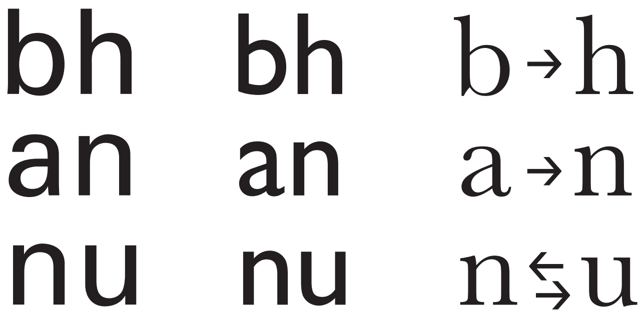


FIGURE 3 *Harris (1973) found that some letter pairs demonstrate a higher misreading in Monotype Baskerville (far right), than the same letter pairs in the two sans serif styles of the investigation (left and middle).*

The findings of Harris suggest that in some letters, serifs can close up the otherwise open counters, which normally would help differentiate one letter from the other. In the case of the 'b' > 'h' confusion, it suggests that the serif on top of the 'b' can trigger a misreading of the letter's lower part, so that it appears to have a gap and to have serifs on both stems. Yet, in the potential confusion of 'i' > 'l' and of 'j' > 'l', the investigation found that the gap between the stem and the dot might be enhanced by the x-height serif, thus improving the legibility of the letter. However, it is possible that the higher legibility of the 'i' in the serif typeface Baskerville might simply be related to the larger distance between the dot and the stem (*figure 4*). The test-typefaces suffer from the problematic issues discussed above. It is possible that the data was influenced by other parameters, such as x-height, letter contrast, weight, and letter skeleton.

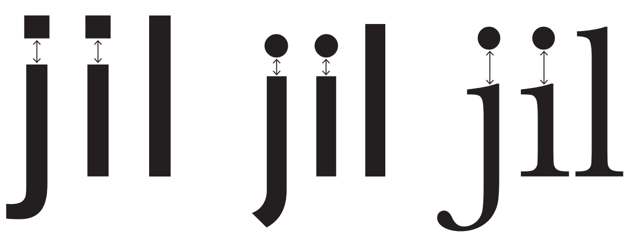


FIGURE 4 *From left, the letters 'j', 'i' and 'l' in the typefaces Univers, Gill Sans, and Monotype Baskerville; note the different distances between the dot and the stem comparing the two sans serif typefaces with the serif style of Baskerville.*

In a legibility investigation of different letter versions within the same typefaces, Beier & Larson (2010) found that a serif on top of the stem of the letter 'i' improves performance at distance viewing compared to the performance of a sans serif ‘i’ of the same typeface. However when the serifs are placed at both the top and the baseline, the identification of the character is just as poor as the version with no serifs at all (*figure 5*).

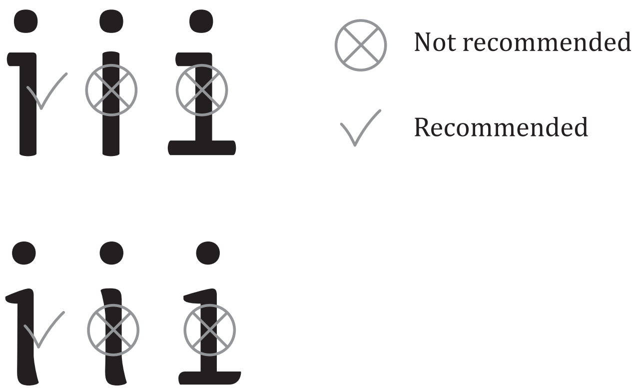


FIGURE 5 *Two of the typefaces tested by Beier & Larson (2010) (top row Ovink, lower row Spencer). Three versions of the letter ‘i’ were tested in each of the typefaces*

The study by Morris and his colleagues (2002) compared the letters as a whole without looking at individual letters; Harris (1973) looked at individual letters, yet applied the problematic method of comparing different typefaces; Beier & Larson (2010) compared different design versions of letters within the same typefaces, yet only examined the effect of serifs on the letter ‘i’.

This study looks into the legibility of the eight letters ‘j’, ‘i’, ‘l’, ‘b’, ‘h’, ‘n’, ‘u’, and ‘a’ identified by Harris (1973) as being influenced by serifs in either a positive or negative direction. The investigation looks at the confusions between these letters both with and without serifs. Following the analysis by Ole Lund (1999) who found no valid conclusion in favour of either serifs or sans serif typefaces, no overall legibility difference between the serif and the sans serif typeface styles is expected. Instead, we wish to investigate whether serifs at the vertical extremes have a positive influence on legibility (Unger 2007; Frutiger n.d.). Therefore the letters l’, ‘b’, ‘h’, ‘n’, ‘u’, all having serifs at their vertical extremes, are compared in their serif and sans serif styles (*figure 6*). It is further expected that the results of the letter confusion comparison will show that some placements of serifs facilitate high legibility and others do not. It is interesting to determine whether the findings are similar to those reported by Harris (1973), or whether a better control of the variables produces a different result.

Harris 1973	jil bh nua
Unger 2007 Frutiger n.a.	l bh nu

FIGURE 6 *The letter group selected for investigation is based on the findings of Harris (1973). Guided by the ideas put forward by Unger (2007) and Frutiger (n.a), a subset of these letters is analysed separately (see Figure 8).*

THE INVESTIGATION

The investigation focuses on the legibility of the letters viewed one at a time. A fluent reader reads a word in a parallel process of letter recognition and word activation (Paap et al., 1982; McClelland & Rumelhart, 1981). This means that by isolating the letter level for investigation, we are only looking at part of the reading mechanism. Yet, if we were to study the individual letter within words, this could cause problems. We know that the different letters within the alphabet vary in legibility (Tinker, 1964). Following this line of thought, it is likely that different neighboring letters also have different levels of influence on each other. Furthermore, the space between the letters plays a central role in reading at a distance; it is possible that the crowding phenomenon (Liu & Arditi, 2001; Hess et al., 2000) where letters blend with each other has a stronger influence in one typeface than another. Consequently, to minimize the influence from other variables in the investigation, we have studied the letters in isolation. The shortcoming of this method is that we will only learn about letters in isolation, so when the same letters are placed within words, we will have no knowledge of how they interact with each other.

There are a number of methods available for testing legibility (for a recent review, see Beier 2012). When focus is on the single letter, the measurement of short exposure and of distance viewing are among the most used. Both of these can be tested in the centre of vision and in parafoveal areas of vision. The findings from the different approaches can, however, differ. One character may have a higher legibility compared to other characters when viewed with a short exposure in the parafoveal area, than the same character viewed from a distance in the centre of vision (for examples see: Beier & Larson 2010). Consequently, to produce findings that are directly applicable to design work, it is important to define the specific reading situation that is under investigation. The focus of the present investigation is the role serifs play when viewed from straight ahead at a distance.

METHOD

In this study, the letters used as test stimuli are designed by the first author to control the variation between the objects of study (the two typeface styles). A sans serif typeface designed for high distance legibility is based on the regular weight of the typeface Ovink (from here on referred to as OvinkSans). For a serif comparison, a new slab serif typeface style named OvinkSerif was designed. The single difference between the letters of the two typeface styles is the added serifs to the typeface OvinkSerif. All other aspects are identical across the two styles (*figure 7*).

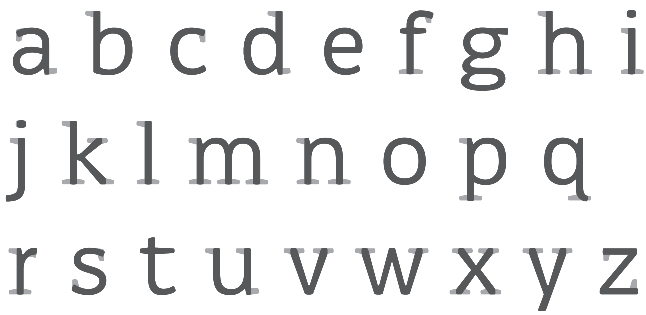


FIGURE 7 *The two typeface styles of OvinkSans and OvinkSerif superimposed.*

The reason for creating a slab serif, rather than a traditional serif typeface, relates to stroke contrast and demonstrates the manner in which design knowledge informs the choice of suitable material for testing. Traditional serifs have a higher stroke contrast than sans serif typefaces. If the serif typeface were of a traditional nature instead of slab serif, both the sans serif and the serif style would have needed a larger stroke contrast than that of OvinkSans. Large stroke contrast in sans serif typefaces is unusual and would therefore have resulted in unnatural test material. A slab serif type was chosen because it can have the same stroke contrast as sans serif type.

It is likely that the performance difference between two conditions that have only one variable that differs is small compared with conditions of test material that have several stylistic features that vary simultaneously (as in much previous research). In the isolation of properties, there is a risk that the difference between test materials is too small for a given method to detect. Yet, this should never be an argument for having several stylistic features simultaneously at work; instead, the sensitivity of the measure needs to be increased (see Poulton, 1965). This was attempted by adjusting the difficulty of the task according to each participant's threshold and increasing uncertainty by mixing styles and including non-target letters (see below). The detailed design of the study drew upon principles from scientific research methods.

A total of 13 participants were recruited from the staff and student community of the Royal Danish Academy of Fine Arts, and they ranged in age from 20 to 61, with an average age of 30. They all reported having normal or adjusted to normal eyesight, and were compensated with a gratuity of 300 Danish kroner.

The study applied a distance threshold method to determine where to position the participant. To locate the region of each participant’s threshold, a pre test of 10 characters presented in a random order was carried out. The letters ‘d’, ’m’, ’p’, ’q’, and ’w’ were shown twice, one time in each typeface style. This choice of letters was based on the analysis of Tinker (1964), who found a fairly consistent trend in these being the most legible letters of the alphabet. By defining each participant’s distance from the stimulus based on this selection, we thereby facilitated misreadings in letters with an assumed lower legibility level. This aspect of the experimental design aimed to produce an appropriate level of performance which would reveal differences in typeface style.

In the pretest, the participant initially stood 10 metres from the stimulus, which was placed 1.7 metres above ground. The participant was then asked to move closer in steps of 0.5 metres until the stimulus was correctly identified. The average distance of the 10 presentations was considered to be the individual's distance. This was the distance at which each participant was positioned throughout the main study.

For both styles of typeface tested, each of the letters ‘j’, ‘i’, ‘l’, ‘b’, ‘h’, ‘n’, ‘u’, and ‘a’, were exposed five times to each participant. The remaining 18 letters of the lowercase alphabet (non-target letters) were exposed once for each typeface style to ensure that participants were presented with the whole alphabet. In this way, participant's responses were not constrained to eight letters which might have increased correct guesses. This resulted in a total of 116 exposures. To familiarize participants with the procedure, the study began with a practice set of 4 presentations. The letters of the two typeface styles were mixed, and each participant was presented with the letters in a different random order. Participants were informed that all material would be lowercase letters and that all the letters of the

English alphabet would be shown with different numbers of occurrences. In order to minimize eyestrain, participants were asked to take as many breaks as they felt necessary. Each letter was presented on white paper at a point size of 22 (with an x-height of 4 mm). The ambient room light was typical for an office environment.

RESULTS

The average number of correct identifications of the letters ‘j’, ‘i’, ‘l’, ‘b’, ‘h’, ‘n’, ‘u’, and ‘a’, across participants is shown in Table 1 for each typeface. This data illustrates which letters are more accurately identified across styles but does not reveal confusions or take account of systematic response biases. For example, if a participant has a tendency to respond 'i', they are likely to get a high correct identification of the letter 'i'. However, their false alarms (i.e. saying 'i' when presented with a different letter such as 'l') have not been taken into account.

	j	i	l	b	h	n	u	a
OVINKSANS	2.31 (0.49)	3.92 (0.40)	1.00 (0.36)	2.38 (0.46)	2.46 (0.35)	2.92 (0.33)	2.08 (0.45)	1.54 (0.42)
OVINKSERIF	2.15 (0.36)	3.08 (0.43)	2.08 (0.45)	2.62 (0.40)	3.00 (0.45)	3.69 (0.31)	2.46 (0.51)	1.62 (0.43)

TABLE 1 *The means (out of a possible 5) of the correct identifications of the 8 letters under investigation, in each typeface style. Standard errors of the means are in parentheses.*

By calculating p(A) using hit and false alarm rates, a measure of sensitivity is obtained which is free of response bias (McNicol, 1972, 113). A value of 0.5 indicates chance performance and 1 equals perfect performance. Table 2 shows the values of p(A) for each letter in each typeface style with the false alarm based on the number of times that any of the other 25 letters is identified as that letter. This represents a more sophisticated analysis of the data than number correct, providing a more accurate account of letter recognition.

Analysis of the data in Table 2 was carried out correcting degrees of freedom for sphericity using the Greenhouse-Geisser estimate of sphericity ($\sum = 0.52$). This shows an effect of letter ($F(7, 84) = 4.17$, $p = 0.0073$), as would be predicted from Tinker (1964). There is no overall difference between the OvinkSans and OvinkSerif ($F(1,12) = 0.88$, $p = 0.36$).

	j	i	l	b	h	n	u	a
OVINKSANS	0.73 (0.05)	0.84 (0.04)	0.61 (0.04)	0.73 (0.05)	0.74 (0.03)	0.78 (0.03)	0.70 (0.05)	0.65 (0.04)
OVINKSERIF	0.71 (0.03)	0.77 (0.04)	0.67 (0.04)	0.76 (0.03)	0.79 (0.05)	0.86 (0.03)	0.75 (0.05)	0.66 (0.04)

TABLE 2

The mean values of p(A) for each letter in each typeface. Standard errors in parentheses.

To test whether serifs at the vertical extremes influence legibility, a further ANOVA was carried out on a subset of the data. The letter group with serifs at the vertical extremes 'l', 'b', 'h', 'n', 'u' show an effect of typeface style ($F(1, 12) = 4.78$, $p = 0.049$) with OvinkSerif letters more reliably identified than OvinkSans (*figure 8*). Again there are differences among letters ($F(4, 48) = 4.54$, $p = 0.0064$, $\sum = 0.83$), with 'n' as the most legible letter of this group.

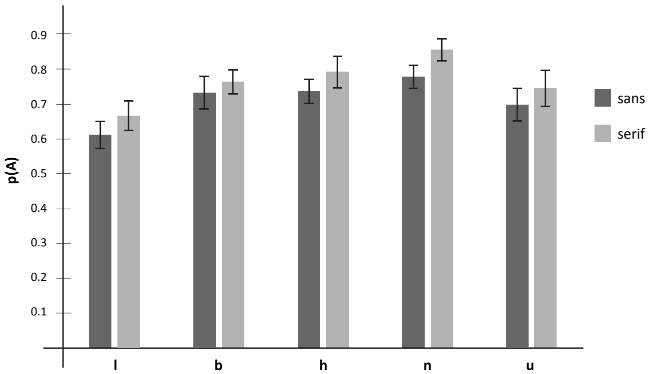


FIGURE 8 *Mean values of p(A) for the 5 test letters that have serifs at the vertical extremes in OvinkSerif, compared with the same 5 letters in OvinkSans. Standard error bars indicate the variability among participants.*

Specific confusions between pairs of letters were also analysed, drawing on the confusions reported by John Harris (1973). These pairs are shown in Figure 9 with the average number of incorrect answers. Given the low number of errors when broken down in this way, it is not possible to calculate p(A) so an ANOVA was carried out on the error scores. There is no main effect of typeface style but a significant difference across the eight pairs ($F(7, 84) = 17.17$, $p < 0.0001$, $\sum = 0.36$); 'l' > 'i' is the most common confusion, resulting in most errors. There is also an interaction between the pairs and typeface style ($F(7, 84) = 4.07$, $p = 0.01$, $\sum = 0.49$). In most pairs, there are more incorrect responses with serifs, but some exceptions. The pairs l > i and i > l and the pairs b > h and h > b appear to account for the interaction. In particular, 'l' is misreported as 'i' more often in Ovink-Sans, whereas 'i' is misreported as 'l' more often with OvinkSerif. 'b' is misreported as 'h' a similar number of times in both typeface styles, whereas 'h' is misreported as 'b' more often in OvinkSerif.

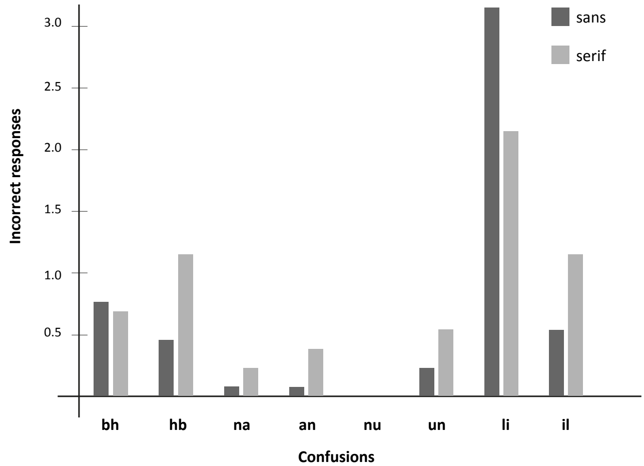


FIGURE 9 *The mean number of incorrect responses between specific letter pairs.*

DISCUSSION

As expected, the overall data for the two typefaces show no statistical difference between OvinkSerif and OvinkSans, indicating that when functioning as whole alphabets, sans serif and serifs typeface styles are equally legible. However, this result obscures the different effect that serifs may have on individual letters, revealed by Harris (1973); when looking at the letters in groups, there are some more subtle findings. The theory put forward by Unger (2007) and Frutiger (n.d.) that serifs at the extremes help facilitate letter recognition, is confirmed in the collective data of the group of letters with serifs at their vertical extremes (‘l’, ‘b’, ‘h’, ‘n’, and ‘u’). The data shows the serif letters were identified significantly better than the sans serif letters.

This indicates that serifs at the vertical extremes do, in fact, facilitate a better definition of the stroke and through this enhance the recognition of letter features.

However, this is not the case in the letters ‘i’ and ‘j’ where the x-height serif — because of the dot — is not at the letters' vertical extreme.

Harris (1973) found a lower confusion in the serif typeface compared with the sans serif typefaces, when combining the ‘i’ > 'l’ and ‘j’ > 'l’ confusions. The present study did not confirm this. Instead, it unexpectedly found the confusion ‘i’ > ‘l’ to be higher when presented in the serif style, and that the reverse confusion ‘l’ > 'i' was higher in the sans serif style. The findings indicate that serifs do not improve legibility on the letter ‘i’, however, play a central role on the letter ‘l’. The idea put forward by Spiekermann (2007) that serifs on ‘i’, ‘l’, ‘I’, and ‘1’ help avoid confusions in the group, is consequently not supported in relation to ‘i’, but confirmed in relation to ‘l’. Combining the findings of the present study with the findings of Beier & Larson (2010), who also looked into different versions of ‘i’ within the typeface Ovink, the results suggest that when serifs are placed at the top and bottom of the stem, they take away the narrow ‘i’ness of the character, and hence result in lower legibility.

Harris (1973) further found the confusion of ‘b’ > ‘h’ to be higher in the serif typeface compared to the sans serif typefaces. This was not confirmed by the present study: there were similar levels of confusion of ‘b’ > ‘h’ for serif and sans serif. However, the reverse confusion (‘h’ > ‘b') was higher in the serif version compared to the sans serif version. This suggests that serifs on the counter of ‘h’ result in the two stems appearing too close, causing the letter to look like a ‘b’.

In the design of new typefaces and logotypes, it is therefore recommended that serifs should be placed at the extremes of those lowercase letters that have vertical stems without a dot on top; that the ‘i’ should be designed without serifs at both ends of the stem; and that serifs should be removed from the counter of ‘h’ (*figure 10*).

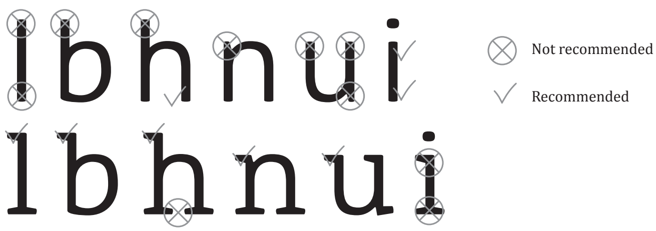


FIGURE 10 *It is recommended that serifs should be placed at the vertical extremes, that ‘i’ has no serifs, and that serifs should be removed from the counter of ‘h’, and most likely also from the counters of ‘n’ and ‘m’.*

We believe that an investigation such as this is far more useful for designers than investigations that, for example, compare the speed of reading Helvetica to the speed of reading Times. The result of such a comparison is only valid to the designer choosing between Helvetica and Times, and is irrelevant in any other situation. It does not enlighten us as to the legibility of the individual features of the typeface style and why participants read the letters the way they do, or to help define the features that can improve the speed of reading in general. Studies that follow the principle of isolating typographical variables contribute more to the field of design. The research not only informs us as to whether one typeface style is more legible than another typeface, but presents findings that help identify the role that different letter features play in a given context.

CONCLUSION

The present investigation found no difference in the distance legibility of sans and serif typeface styles when studying the collective results of the individual letters ‘j’, ‘i’, ‘l’, ‘b’, ‘h’, ‘n’, ‘u’, and ‘a’. However, looking at a group of letters with serifs at the vertical extremes, the experiment demonstrated higher distance legibility when these letters have serifs on the stems; this demonstrates that serifs on such locations play a central role in facilitating high distance legibility. The data further indicates that serifs on the counter of ‘h’ and on both ends of the stem of ‘i’ cause a higher misreading for ‘b’ and ‘l’ respectively. The findings as a whole suggest that serifs should not always be applied in the conventional fashion as seen in traditional Old Style and Didone typefaces; instead they have an ability to facilitate higher distance legibility if they are placed in a semi-serif fashion where they are positioned at relevant stroke endings.

The results were reached through an interdisciplinary collaboration between researchers from the fields of science and design. By designing typefaces that allow for isolation of a single typographical variable, and through the creation of an experiment that adopts a scientific approach, this study achieves a high level of internal validity. By adopting a research question that is relevant to the design community, the research further provides a high level of external validity. The detailed and specific nature of these results can inform character design by identifying features that influence letter recognition. This paper has aimed to provide a foundation for future research through espousing the benefit of integrating scientific and design expertise,and by providing an example of the application of this approach. Further work could build upon these findings to pursue the role of serifs in more detail or to investigate other typographic stylistic features of typefaces (e.g. weight or stroke contrast) in relation to particular reading situations.

There is no doubt that a sceptical approach to the field of experimental legibility investigations has its merits; the many studies based on poor typographical test material has muddied the waters and overshadowed a number of valid results. Reading is a complicated process; consequently, so is legibility. The most successful experimental investigations of typefaces are the ones that manage to simplify the parameters by isolating one variable for investigation, and through that provide data that can inform the practice of design.

REFERENCES

Beier, Sofie. 2012. *Reading Letters: designing for legibility*, BIS Publishers.

Beier, Sofie & Larson, Kevin. 2010. Design Improvements for Frequently Misrecognized Letters. *Information Design Journal*, 18.2, 118-137.

Berry, John. D. 2006. Warm modernism in profile, in: J.D. Berry (ed.) *Dot-font: Talking About Fonts.* New York: Mark Batty Publisher, 65-68.

Bias, Randolph. G., Larson, Kevin., Huang, Sheng-Cheng., Aumer-Ryan, Paul R., & Montesclaros, Chris. 2010. An exploratory study of visual and psychological correlates for on-screen sub-pixel rendered text. *Journal of the American Society of Information Science and Technology,* 61.4, 745-757.

Cross, Nigel. (2001). Designerly Ways of Knowing: Design Discipline versus Design Science. *Design Issues.* 17.3, 49-55.

De Lange, R.W., Esterhuizen, H.L., & Beatty, D. 1993. Performance differences between Times and Helvetica in a reading task. *Electronic Publishing,* 6.3, 241-248.

Dyson, Mary C. 2011. Do Designers Show Categorical Perception of Typefaces?, *Visible Language,* 45.3., 193-220.

Dyson, Mary C. and Kipping, G. J. 1998. The effects of line length and method of movement on patterns of reading from screen. *Visible Language* 32.2, 150-181.

Fiset, Daniel., Blais, Caroline., Éthier-Majcher, Catherine., Arguin, Martin., Bub, Daniel., & Gosselin, Frédéric. 2008. Features for Identification of Uppercase and Lowercase Letters. *Psychological Science.* 19.11, 1161-1168.

Fox, Doug., Chaparro, Barbara S., & Merkle, Ed. 2007. Examining Legibility of the Letter “e” and Number “0” Using Classification tree Analysis. *Usability News,* 9(2), http://www.surl.org/usabilitynews/92/legibility.asp (Accessed February 1, 2013).

Frutiger, Adrian (n.d.), The History of Linear, Sans Serif Typefaces, http://www.linotype.com/2258-16893/thelossoftheserifspart1.html (Accessed February 15, 2013).

Harris, John. 1973. Confusions in Letter Recognition. *Professional Printer.* 17.2, 29-34.

Hess, R.F., Dakin, S.C. & Kapoor, N. 2000. The Foveal ‘Crowding’ Effect: Physics or Physiology?. *Vision Research.* 40, 365-370.

Heylighen, Ann., Cavallin, Humberto., & Bianchin, Matteo. 2009. Design in Mind. *Design Issues.* 25.1, 94-105.

Larson, Kevin & Sheedy, James E. 2008. Blink: The Stress of Reading. Eye*.* 67. 95.

Liu, Lei. & Arditi, Aries. 2001. How crowding affects letter confusion. *Optometry and Vision Science.* 78.1, 50-55.

Lund, Ole. 1999. Knowledge Construction in Typography: The case of legibility research and the legibility of sans serif typefaces. Thesis PhD, The University of Reading, Department of Typography & Graphic Communication.

Luckiesh, Matthew. & Moss, Frank K. 1942. *Reading as a Visual Task,* New York: D. Van Nostrand Company.

Mansfield, J. S., Legge, G. E., & Bane, M. C. 1996. Psychophysics of reading: XV. Font effects in normal and low vision. *Investigative Ophthalmology & Visual Science,* 37, 1492-1501.

McClelland, James L. & Rumelhart, David E. 1981. An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Account of Basic Findings. *Psychological Review.* 88.5, 375-407.

McNicol, D. 1972 *A Primer of Signal Detection Theory.* London: George Allen & Unwin Ltd.

Morris, Robert A., Aquilante, Kathy., Yager, Dean. & Bigelow, Charles. 2002. P-13: Serifs Slow RSVP Reading at Very Small Sizes, but Don’t Matter at Larger Size. *SID Symposium Digest of Technical Papers.* 33.1, 244-247.

Nelson, Harold G. & Stoltenman, Erik. 2003. *The Design Way: Intentional Change in an Unpredictable World.* Educational Technology Publications.

Paap, K.R., Newsome, S.L., McDonald, J.E., & Schvaneveldt, R.W. (1982) 'An Activation-Verification Model for Letter and Word Recognition: The word superiority effect', *Psychological Review, 89.5*, 573-594

Poulton, E. C. 1965. On increasing the sensitivity of measures of performance. *Ergonomics. 8.1*, 69-76.

Pyke, Richard L. 1926. *Report on the Legibility of Print.* Medical Research Council, London: His Majesty's Stationery Office.

Roethlein, Barbara E. 1912. The relative legibility of different faces of printed types. *American Journal of Psychology. 23.1*, 1-36.

Sheedy, J. .E., Tai, Y.Subbaram, M.,Gowrisankaran, S. & Hayes, J. 2008. ClearType sub-pixel text rendering: Preference, legibility and reading performance. *Displays.* 29.2, 138-151.

Soleimani, Hassan, & Mohammadi Elham. 2012. The effect of text typographical features on legibility, comprehension, and retrieval of EFL learners. *English Language Teaching.* 5.8, 207-216.

Spiekermann, Erik. 2007. How does the serif on a sans-serif i increase legibility/readability? (Comment in the discussion) http://typophile.com/node/38217 (Accessed February 1, 2013).

Tinker, Miles A. 1964. *Legibility of Print.* U.S.A.: Iowa State University Press.

Typophile. 2008a. Legibility/Aesthetic — Improving the reading experience. http://www.typophile.com/node/50834 (Accessed February 1, 2013).

Typophile. 2008b. Know any good legibility tests for typefaces?. http://www.typophile.com/node/41365 (Accessed February 1, 2013).

Unger, Gerard. 2007. *While You're Reading,* New York: Mark Batty Publisher.

Waller, Robert. 2007. Comparing typefaces for airport signs. *Information Design Journal,* 15.1, 1-15.

Wrolstad, Merald E. 1969. Letterform Research Needs Definition and Direction: A Report from the Editor. *Journal of Typographic Research.* 3.2, 115-126.

ABOUT THE AUTHORS

Dr. Sofie Beier is a type designer and assistant professor employed by the School of Design under The Royal Danish Academy of Fine Arts. She holds a PhD from the Royal College of Art in London, and is the author of the book *Reading Letters: Designing for Legibility*. Her research focuses on improving the reading experience by gaining a better understanding of how different typefaces and letter shapes can influence the way we read.

www.sofiebeier.dk

Dr. Mary Dyson is Senior Lecturer in the Dept of Typography & Graphic Communication, University of Reading, UK. She studied experimental psychology leading to a PhD in perception, and now teaches theoretical and empirical approaches to typography and graphic communication. She has supervised many research students over the years on quite diverse topics and continues to enjoy this activity. Her most recent research interests are driven by a desire to bridge the gap between scientists and designers and find commonalities. Her experiments have looked at issues such as how designers perceive typefaces compared to non-designers and how Chinese typefaces are perceived by non-Chinese readers, with the aim of understanding how typefaces are processed in reading.

http://www.reading.ac.uk/typography/about/staff/m-c-dyson.aspx

AUTHOR CONTACT

Dr. Sofie Beier
School of Design
The Royal Danish Academy of Fine Arts
Philip de Langes Allé 10
DK-1435 København K
Denmark
sbe@kadk.dk

Dr. Mary C Dyson
Department of Typography & Graphic Communication
University of Reading
2 Earley Gate
Whiteknights
Reading RG6 6AU
M.C.Dyson@reading.ac.uk

NEW VISIBLE LANGUAGE ADVISORY BOARD MEMBERS

This year Visible Language has transitioned from editor Sharon Poggenpohl to Mike Zender and settled into its new home at the University of Cincinnati. With that transition has come a renewed focus on visual communication research and with that the addition of several new Advisory Board members. Short biographical sketches for three of our new board members follow here. More will follow as space permits.

MARY C DYSON

Mary C Dyson is an Associate Professor in the Department of Typography & Graphic Communication, University of Reading, UK. She studied experimental psychology leading to a PhD in perception, before switching discipline to teach theoretical and empirical approaches to typography and graphic communication. The focus of her teaching and research has been how users interact with documents, and at a more specific level, how typefaces are processed when reading and when designing. She has done experiments with screen-based material and published work on reading and interacting on screen alongside reviews of other research in this area. She has also supervised many research students on topics relating to her own research, but also more broadly within the field. This experience has developed her teaching of research methods.

Her research interests are driven by a desire to bridge the gap between scientists and designers and find commonalities. She is therefore committed to interdisciplinary research and enjoys collaborating with colleagues from various disciplines to explore areas of common ground.

Recent work has drawn on her PhD in perception and has looked at the perception of typefaces by typography students, in comparison to non-designers. These studies draw on examples of research into other areas of perception, both visual and auditory, i.e. the perception of faces, music, and speech, which suggest avenues to explore in relation to how we perceive visual forms. In particular, this approach stimulates ideas concerning particular methods of investigation and seeks to develop novel experiments within the field of typography.

JORGE FRASCARA

Jorge Frascara is Professor Emeritus (University of Alberta), Fellow of the Society of Graphic Designers of Canada and of the Society for the Science of Design of Japan, Advisor to the Doctoral Program at the University IUAV of Venice, and Adjunct Professor at the Universidad de las Americas Puebla. He was an advisor to the ISO and to the Canadian Standards Council on graphic symbols. He has been President of Icograda and Chairman of the Department of Art and Design at the University of Alberta.

He is the author of Communication Design (2005); and User-Centred Graphic Design (1997); and the editor of Designing Effective Communications (2006); Design and the Social Sciences (2002); Graphic Design, World Views (1990); and the ISO Technical Report 7239, Design and Application of Public Information Symbols (1984). He has also published four books in Spanish and more than 50 articles internationally. He is an advisor for four design journals, and has received honors and awards from a wide range of organizations.

Frascara has lived and worked worked in Argentina, Canada, Guatemala, England, Italy, and Mexico, has been a guest lecturer in 26 countries, and during his 31 years in Canada he was a consultant for different departments of the Federal Government, the Province Alberta, Telus Canada, the Mission Possible Coalition (traffic safety), the Alberta Drug Utilization Program, and other organizations. In Italy he worked for the Health Services, and for traffic safety. He now lives in Cholula, Mexico, and runs an information design and social communications consultancy with his wife Guillermina Noël.

KEN FRIEDMAN

Ken Friedman is University Distinguished Professor and Dean of the Faculty of Design at Swinburne University of Technology in Melbourne, Australia. He works at the intersection of three fields: design, management, and art. Friedman works with theory construction and research methodology for design. He also works with design process and design thinking as tools for value creation and economic innovation. He is active in developing international research networks and conferences for the design research community.

Friedman is an editor of the journals *Artifact* and the *Journal of Design Research*, and a member of the editorial board of such journals as *Design Studies, Design and Culture*, the *International Journal of Design*, and *Visible Language*.

Ken is also a practicing artist and designer active in the international laboratory known as Fluxus. He had his first solo exhibition in New York in 1966. His work is represented in major museums and galleries around the world, including the Museum of Modern Art and the Guggenheim Museum in New York, the Tate Modern in London, the Hood Museum of Art at Dartmouth College, and Stadtsgalerie Stuttgart.

STAN RUECKER

Dr. Stan Ruecker is an Associate Professor with current research interests in the areas of humanities visualization, the future of reading, and information design. He came to ID from the University of Alberta's interdisciplinary Humanities Computing program where he was also an Associate Professor, supervising graduate students and leading seminars on experimental interface design, knowledge management and analysis, research methods, and interdisciplinary research project management. His students have gone on to work with major research and development projects in fields ranging from medical imaging to oilfield decision support.

He is a major grant holder, and his research teams have presented their findings at over a hundred international conferences in design, computing science, educational technology, literature, communication technology, library and information studies, and humanities computing. He was the principal investigator of the SSHRC SRG Humanities Visualization team, and currently leads the interface design unit of the SSHRC MCRI Implementing New Knowledge Environments (INKE) project.

His work to date has focused on developing prototypes to support the hermeneutic or interpretive process, and he has published extensively on information design, experimental interface design, and interdisciplinary research project management. His book *Visual Interface Design for Digital Cultural Heritage*, co-authored by Milena Radzikowska and Stéfan Sinclair, was released in 2011 by Ashgate Press.

He holds an interdisciplinary PhD in Humanities Computing from University of Alberta, an MDes from the same, an MA in English literature from University of Toronto, and advanced undergraduate degrees in English literature and computer science from University of Regina. He is an Adjunct Professor at the University of Alberta's School of Library and Information Studies, Department of English and Film Studies, and Humanities Computing Program, and in the University of Victoria's Faculty of Humanities.

JOURNAL INFORMATION

Visible Language is an academic journal focused on research in visual communication. We invite articles from all disciplines that concern visual communication that would be of interest to visual communication designers.

READERSHIP

Visible Language, an academic journal focused on research in visual communiction, seeks to advance research and scholarship for two types of readers: academics and professionals. The academic is motivated to consume knowledge in order to advance knowledge thorough research and teaching. The professional is motivated to consume and apply knowledge to improve practice. *Visible Language* seeks to be highly academic without being inaccessible. To the extent possible given your topic, *Visible Language* seeks articles written to be accessible to both our reader types. Anyone interested may request a copy of our editorial guidelines for authors.

EDITORIAL CORRESPONDENCE

Article concepts, manuscripts, inquiries about research and other contributions to the journal should be addressed to the editor. We encourage article concepts written as an extended abstract of 1 to 2 pages single-spaced. We will offer prompt feedback on article concepts with our initial opinion on their suitability for the journal. Manuscripts accepted for peer review will receive a summary response of questions or comments within three weeks. Letters to the editor are welcome. Your response — and the author's reply — will not be published without your permission and your approval of any editing. If you are interested in submitting an article to the journal and would like a copy of our Notes on the Preparation of a Manuscript, please obtain it from the journal's website at visiblelanguagejournal.com

Editorial correspondence should be addressed to:

Mike Zender
Editor, *Visible Language*
College of Design, Architecture, Art, and Planning
School of Design
University of Cincinnati
PO Box 210016
Cincinnati, OH 45221-0016
mike.zender@uc.edu

If you are interested in serving as guest editor for a special issue devoted to your specific research interest, write to the editor, outlining the general ideas you have in mind and listing a half dozen or so topics and possible authors. If you would rather discuss the idea first, call the editor at: 513.556.1072

BUSINESS CORRESPONDENCE

Subscriptions, advertising and related matters should be addressed to:

Visible Language
Sheri Conttingim
Office of Business Affairs
College of Design, Architecture, Art, and Planning
University of Cincinnati
PO Box 210016
Cincinnati, OH 45221-0016
513.556.4377
sheri.cottingim@uc.edu